(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 October 2003 (02.10.2003)

PCT

(10) International Publication Number
WO 03/081857 A1

(51) International Patent Classification[7]: H04L 12/56, 12/46

(21) International Application Number: PCT/US03/06641

(22) International Filing Date: 4 March 2003 (04.03.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/103,470 20 March 2002 (20.03.2002) US

(71) Applicant: **INTEL CORPORATION** [—/US]; 2200 Mission College Boulevard, Santa Clara, CA 95052 (US).

(72) Inventors: **NAVADA, Muraleedhara**; 3707 Poinciana Drive, #M1101, Santa Clara, CA 95051 (US). **KURUPATI, Sreenath**; 2016-3 Klamath Avenue, Santa Clara, CA 95051 (US).

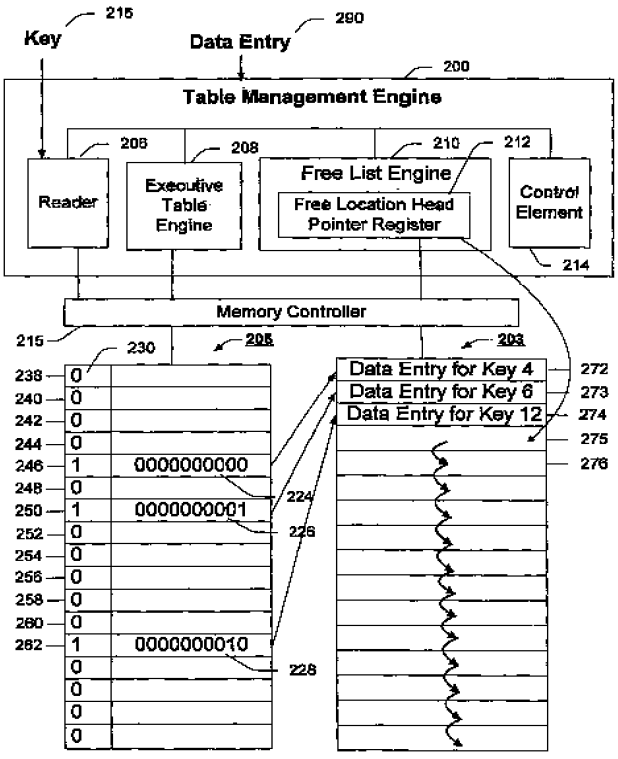
(74) Agents: **MALLIE, Michael, J.** et al.; Blakely, Sokoloff, Taylor & Zafman LLP, 7th floor, 12400 Wilshire Boulevard, Los Angeles, CA 90025 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: VLAN TABLE MANAGEMENT FOR MEMORY EFFICIENT LOOKUPS AND INSERTS IN HARDWARE-BASED PACKET SWITCHES



(57) Abstract: Described herein is a method and apparatus for memory efficient fast VLAN lookups and inserts in hardware-based packet switches. A table management engine (TME) is introduced into the switches to accelerate date searches and management. The TME manages a memory which is devided into two tables, one containing the data entries ("data entry table" 203), one containing the pointers and validity bits ("pointer table" 205). The TME comprises a reader/hasher (206), an executive table engine (208), a free list engine (210) and a control element (214). The reader/hasher obtains content from a key, which can be a VLAN rule. The executive table engine uses the kex content to address a pointer location in memory to retrieve the data. The executive table engine is also responsible for writing /deleting data into the memory and for pointers management.

WO 03/081857 A1

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

VLAN TABLE MANAGEMENT SYSTEM FOR MEMORY EFFICIENT LOOKUPS AND INSERTS IN HARDWARE-BASED PACKET SWITCHES

TECHNICAL FIELD

[0001] The present invention generally relates to the field of data processing and specifically to memory efficient fast VLAN lookups and inserts in hardware-based packet switches.

BACKGROUND

[0002] As computing and networking devices become faster, the requirement for speed in the management of data tables challenges conventional approaches. The speed of a key search, for example, in which data associated with a key must be found in a table quickly, has become a critical issue, and sometimes a bottleneck, in many devices and applications. A key may be any piece of data used as an index or search criterion for finding additional data, but in a networking context, keys are typically Internet protocol (IP) addresses, media access control (MAC) addresses, virtual local area network (VLAN) tags, and other network identifiers.

[0003] Solutions that accelerate key search speed sometimes depend on the characteristics of the memory used to store the data table being searched. Random access memory (RAM) stores data at a particular location denoted by an address. When the address is supplied to the RAM, the RAM returns the data stored there. To find the correct address, however, either an index of all the keys needs to be sorted and searched for an address associated with the key or all the associated data entries must be searched for a representation of the key and its associated RAM address. There are many algorithms that seek to shorten the search time for an address associated with a key.

[0004] One type of hardware memory, content addressable memory (CAM), accelerates the search for a stored data item by retrieving the data based on the content of the data itself, rather than on its address in memory. When data is supplied to a CAM, the CAM directly returns an address where the associated data is found. For many applications, CAM provides better performance than conventional memory search algorithms by comparing desired information against an entire list of stored data entries simultaneously. Hence, CAM is used in applications in which search time is an important issue and must be constrained to very short durations.

[0005] Unfortunately, both discrete hardware and integrated circuit CAM implementations can be relatively expensive both in chip area requirements and/or design complexity. In some applications a direct-mapped cache could be used as a substitute for a CAM, but the fully associative characteristic of a CAM--where a data entry can be placed anywhere in the data structure--is lost and undesirable characteristics such as data collisions and unused memory locations are introduced.

BRIEF DESCRIPTION OF THE DRAWINGS

[0006] The present invention is illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings in which like reference numerals refer to similar elements and in which:

[0007] **Fig. 1** is a block diagram of an example computing network system employing example table management engines (TMEs), according to a networking embodiment of the invention;

[0008] **Fig. 2** is a block diagram of an example TME, according to one embodiment of the invention;

[0009] **Fig. 3** is a graphical representation of an example memory being indexed and/or accessed according to the content of a key, in accordance with one embodiment of the invention;

[0010] **Fig. 4** is a block diagram of an example executive table engine of Fig. 2;

[0011] **Fig. 5** is a graphical representation of an example linked-list of free data entry memory locations, in accordance with a data insertion embodiment of the invention;

[0012] **Fig. 6** is a graphical representation of an example linked-list of free data entry memory locations, in accordance with a data deletion embodiment of the invention;

[0013] **Fig. 7** is a flowchart of an example data entry retrieval method, according to one embodiment of the invention;

[0014] **Fig. 8** is a flowchart of an example key insertion method, according to one embodiment of the invention;

[0015] **Fig. 9** is a flowchart of an example data entry deletion method, according to one embodiment of the invention; and

[0016] **Fig. 10** is a graphical representation of an example storage medium comprising content which, when executed, causes an accessing machine to implement one or more embodiments, aspects, and/or methods of a TME.

DETAILED DESCRIPTION

[0017] The present invention is generally directed to a method and apparatus for memory efficient fast VLAN lookups and inserts in hardware-based packet switches.

[0018] In accordance with the teachings of the present invention, a table management engine (TME) is introduced to accelerate data table searches and management. Because so many computing and networking devices store and retrieve data, a TME can improve

the data lookup speed and management performance of many types of tables, lists, and databases. For example, in the data communications field, the TME can speed up devices and applications that use address tables, translation tables, filter tables, and/or VLAN rule tables.

Example Context for Implementing Aspects of the Invention

[0019] **Fig. 1** is a block diagram of an example computing network 100 in which TMEs 102, 104, 106 accelerate the performance of data packet switching and routing. The illustrated example network 100 includes a router 112, two switches 108, 110 and six computing hosts 120-130 communicatively coupled as illustrated. The example computing network 100 is logically divided into VLAN A 114, VLAN B 116, and VLAN C 118. Example network devices, such as the illustrated switches 108, 110 and router 112 typically use internal tables to associate a source address for each data packet received with destination information stored in the tables.

[0020] Using VLAN rules as an example of destination information, each VLAN rule may be one or more port and/or destination addresses and/or other packet directing information. Accordingly, each data packet is sent to a proper hardware address and/or an IP address (depending on the device) as directed by a particular VLAN rule. Although the computing network 100 is depicted as one environment affording a context in which TMEs 102, 104, 106 could be implemented, many other environments and uses are possible.

[0021] As a computing network 100 routes and directs data packets, one or more TMEs 102, 104, 106 can be situated in the various components that participate in networking, such as in the illustrated router 112 and two switches 108, 110. TMEs 102, 104, 106 could be implemented in hosts 120-130 and clients as well. In fact, TMEs 102, 104, 106 can be

used in any part of the computing network 100 where data lookups occur and/or a table of information is kept.

[0022] TMEs 102, 104, 106 can enhance the performance and management of the IP address, port address, and hardware address tables in a computing network 100. When a data packet is received in a device that forwards the data packet using the information in the data packet to decide the forwarding, a TME (102, 104, 106) uses content from the data packet to index directly into a table that stores (or is able to store) a data entry corresponding to the data packet. The data entry is used for other data packets that possess the same content or that can be read and/or hashed to supply the same content.

[0023] The data entry contains the forwarding information. In other words, content from the data packet is used directly as a memory address for finding the key's associated data entry without further performing a search through a list of keys to find a memory address for the key. In simplest terms, the key content is substantially used as a memory address. TMEs 102, 104, 106 perform this function of addressing memory by content without the addition of known CAM hardware elements.

[0024] In the illustrated computing network 100, a data packet from a first host 120 in VLAN A 114 is received at a first switch 108 having a TME 104. The TME 104 reads and/or hashes the data packet for a content, such as the VLAN tag of VLAN A 114. This content corresponds to a location in a table that is set up so that the table locations correspond to various contents obtainable (readable and/or hashable) from data packets. Thus, the content directly provides an address for indexing into the table and obtaining the destination information (or a pointer to the destination information) for the data packet without using a search algorithm to find a memory address corresponding to the content. The destination information may be a VLAN rule having port and address information. The addressing of the table by content is performed without known CAM hardware

elements. Using the destination information in the data entry for the key, the data packet is forwarded to the second switch 110 also having a TME 106. The TME 106 addresses its data entry table by content as described above, and directs the data packet to the router 112. The router also possesses a TME 102 which functions as described above. Thus, the data packet is directed and routed through various network devices to reach its final destination, the second host 122 in VLAN A 114. The TMEs 102, 104, 106 provide faster data packet directing than known non-CAM methods, and in the illustrated example, provide the faster VLAN lookups while making more efficient use of memory than known methods that require a search and/or lookup algorithm to find a memory address for each key.

Example Architecture

[0025] **Fig. 2** is a block diagram of an example TME 200, according to one implementation of the invention, for accelerating the performance of data tables 203, 205 in one or more memories ("in memory"). An overview of example components is provided to introduce an example architecture. In this embodiment, the TME 200 includes a reader/hasher ("reader") 206, a executive table engine 208, a free list engine 210, a free location head pointer register 212, and a control element 214 communicatively coupled as illustrated. The TME 200 is coupled to a memory controller 215, which is coupled to the tables 203, 205 in memory. More than one memory controller 215 could be used. In this implementation, a first table 203 contains data entries ("data entry table" 203), and a second table 205 contains pointers and validity bits ("pointer table" 205). It should be noted that the table(s) 203, 205 included in or used by a TME 200 can be correctly

referred to as either a single table or multiple tables, since they are relational and may be split in different ways and in various different memories or memory parts.

[0026] For purposes of explaining aspects of the invention, the first table 203 may also be called the first memory 203 and/or the data entry memory 203. The second table 205 may also be called the second memory 205 and/or the pointer memory 205. Accordingly, a specific location in a table may also be called a memory location. Those skilled in the art will appreciate that these alternate terms for the first table 203 and the second table 205 and specific locations therein are substantially equivalent because tables used in computing and networking devices are commonly implemented in memory. The data entries, pointers, and validity bits in the tables and/or memories may also be stored in different arrangements.

Memory Addressability by Content

[0027] The example TME 200 receives data, such as a key 216 and/or a data entry 290. If the application is a computing network 100, the key 216 may be a hardware address, software address, and/or VLAN tag included in the header of a data packet/datagram, but in other non-IP applications the key 216 may be any data. The content that is obtained from the key 216 is used to directly address and/or index into a table/memory location. "Directly" as used here means that no search algorithm is needed to sort through a list and/or index file containing multiple keys or other data entries to relate a memory address stored in the list/index file to the key 216. For example, as illustrated in **Fig. 3**, an example content value of "7" 302 from a key is substantially the address and/or position of a memory location "7" 304 for the key, thereby providing content addressability to

standard types of memory, such as RAM, and avoiding a search through a list of keys, contents, or records.

[0028] Returning to Fig. 2, the reader 206 obtains a content from the key 216 using all or part of the data in the key 216. That is, in addition to merely reading the key 216 to obtain its content, the reader 206 may also obtain the content by hashing all or part of the key 216. In this regard, the content determined by the reader 206 does not have to be a literal one-to-one translation of data in the key. The only requirement for the reader 206 and for the content obtained is that the same content is always obtainable from a given key 216. In a networking context, this means that data packets yielding the same content will be directed to the same destination(s). In addition to directly reading and/or hashing the key 216 to obtain content, the content may also be obtained by substituting predetermined content for the actual content read by the reader 206, that is, by bit masking, and/or by other methods that yield a reproducible content from a given key 216.

[0029] To further illustrate content addressability according to one aspect of the invention, Fig. 2 shows a pointer 224 for a key with content "4" at the fourth (counting from "0") location 246 of the second table 205, a pointer 226 for a key with content "6" at the sixth location 250 of the second table 205, and a pointer 228 for a key with content "12" at the twelfth location 262 of the second table 205. It should be noted that unlike the *location* of a pointer in a table or in memory, the actual value of the pointer (the address that the pointer is pointing to) does not correspond to the content of the key it is associated with, unless coincidentally. Rather, the pointer's address value is directed to a data entry location somewhere in the first table 203.

[0030] **Fig. 4** is a block diagram of an example executive table engine of Fig. 2. A key indexer 402, a pointer engine 404, a data entry retriever 406, and a data entry

inserter/deleter 408 are coupled as shown. The key indexer 402 receives a key 216 content from the reader 406, and uses the content to index into the second memory 205. At the memory location 246 corresponding to the content of the key 216, the key indexer reads the validity bit stored there to ascertain whether there is a valid data entry associated with the key 216. The key indexer 402 addresses the second memory 205 directly using the content of the key 216, because the content of the key 216 is substantially the needed memory address.

[0031] Using a key content of "4" as an example, the pointer engine 404 establishes a pointer 224 between a newly inserted data entry at memory location 272 and a memory location 246 corresponding to the content "4" of the key 216 in the second memory 205. The pointer 246 established by the pointer engine 404 is the address of the memory location 272 containing the data entry in the first memory 203. During a key deletion operation, the pointer engine 304 also deletes the pointer 224 and sets the associated valid bit to zero.

[0032] In some embodiments, the pointer engine 404 also performs the function of setting the validity bit when a pointer operation is carried out, that is, sets the validity bit to indicate the presence or absence of a pointer and therefore a data entry associated with a key 216.

[0033] The data entry retriever 406 reads the pointer 224 stored at the memory location 246 provided by the key indexer 402, and follows the pointer 224 to the memory location 272 in the first memory 203, returning the stored data entry.

[0034] When the data entry inserter/deleter 408 inserts or deletes a data entry stored in a memory location in the first memory 203 it notifies the pointer engine 404 to add or delete, respectively, the pointer 224 from the second memory 205 and to assign the

validity bit at the memory location 246 to reflect the presence or absence, respectively, of the data entry being added or deleted.

VLAN Embodiment

[0035] In the context of tables used for IP data communications, an IP address from the header of a data packet/datagram ("packet") is often used as the key 216 to find associated information about the packet, such as a destination port number, the sender's access privileges and location on a network, or applicable VLAN rules. In one example embodiment, the example TME 200 is used as a VLAN rule table or to manage a VLAN rule table.

[0036] For example, in Ethernet switches, VLAN rules need to be stored and looked up using the VLAN tags of incoming data packets. A certain number of VLAN rule entries, for example 1K entries, need to be stored and looked up for incoming packets.

[0037] The TME 200 can take advantage of the fact that VLAN tags are 12 bits wide to provide the functionality that a VLAN rule and/or address space would have if implemented in a traditional CAM chip, but without the chip area requirements and/or design complexity of CAM hardware. Like a CAM, the TME 200 can utilize all the memory locations in a selected RAM, preventing packet collisions, and emulating the content addressability of a CAM.

[0038] Although there may be memory overhead when implementing a VLAN rule table in RAM, such as a 1K RAM, using the TME 200 as a VLAN rule table and/or VLAN rule table manager is better than using a fixed-size hash structure and/or cache to perform VLAN rule lookups, because the TME 200 provides CAM functionality thus guaranteeing (the example 1K) address space by preventing packet collisions. The latter feature is

critical for chip vendors using the TME 200 who must guarantee that a certain number of entries can be stored.

[0039] Referring to Fig. 2, various size RAMs can be selected for the first memory 203 and the second memory 205 when implementing a VLAN rule table using the TME 200. For the first memory 203, a 1K RAM (or as large a memory capacity as desired) could be used to store the example 1K VLAN rule entries. For the second memory 205, since 4K is the usual maximum number of VLAN rule entries needed in a VLAN rule table, a complete 4K RAM could be used to store a maximum of 4K pointers and associated validity bits.

[0040] For the pointer and the validity bit to be stored in a memory location in the second memory 205, a width of eleven bits is sufficient for a VLAN rule table having a depth of 1K. Each memory location in the pointer table 205 could have one bit allotted for the validity bit, and ten bits allotted for the pointer. Ten-bit pointers 224, 226, 228 are used in the example, because a binary number having ten bits can represent the desired 1K (1024 bits) memory locations for the VLAN rule entries in the data entry table 202. The length of the VLAN rule entries can vary and can be accommodated by selecting a wide- enough memory, for example a 200 bit wide memory for correspondingly wide rule entries. Thus, the TME 200 provides a convenient, high-speed, and memory efficient VLAN rule table with all the advantages that a CAM chip would provide without using a CAM chip.

Free List Engine

[0041] Referring still to Fig. 2, a TME 200 may incorporate a free list engine 210 in some embodiments. The free list engine 210 manages and maintains a list of available ("free") memory locations (e.g., 275-276) in the first memory 203. In one embodiment, the free

location head pointer register 212 is included in the free list engine 210 to point to the first available free memory location 275 in a list of free memory locations 275-276. The first available free memory location 275 is allocated to the next data entry to be inserted, unless some other occupied data entry location 272-274 becomes free first.

[0042] In one embodiment, a linked-list of free memory locations 275-276 is used for storing data entries associated with keys. To illustrate example operations for maintaining the linked-list, **Fig. 5** is a graphical representation of an example linked-list of free memory locations, in accordance with a data insertion aspect of the invention. A first memory 203 for storing data entries associated with keys, a free list engine 210, and a free location head pointer register 212 are example components that participate in maintaining the linked-list.

[0043] In an example data entry insertion, a data entry for key "13" 502 is inserted into the first available memory location 275. The address of the next free memory location 276 is transferred to the free location head pointer register 212 to update the head pointer 504 so that it no longer points to the now occupied memory location 275, but instead points 506 to the next free memory location 276 in the linked-list. The next free memory location 276 now becomes the first available free memory location.

[0044] Like the previous figure, **Fig. 6** is a graphical representation of an example linked-list of free memory locations, in accordance with a data deletion aspect of the invention. The first memory 203 for storing data entries associated with keys, and a free list engine 210, having a free location head pointer register 212 are among the example components that participate in the operation.

[0045] To illustrate example dynamics for maintaining the linked-list of free memory locations during a data entry deletion, consider the state of the free location head pointer register 212 and the linked-list of free memory locations before the data entry deletion.

The free location head pointer register 212 contains the address of the first available free memory location 276, represented by a pointer 604. The data entry deletion then occurs: the data entry for key "12" 602 is deleted from its memory location 274. The address of the first available free memory location 276 is copied from the free location head pointer register 212 into the newly emptied memory location 274, establishing a pointer 603 pointing from the newly emptied memory location 274 to the (former) first available free memory location 276. The address of the newly emptied memory location 274 is copied into the free location head pointer register 212. Thus, the newly emptied memory location 274 becomes the new first available free memory location at the head of the linked-list of free memory locations. The memory location from which the data entry is being deleted always becomes the first available free memory location at the head of the linked-list of free memory locations, in this embodiment.

[0046] Having described operations that can be performed by a TME 200, it will be appreciated by those having ordinary skill in the art that variations in the architecture of a TME 200 are allowable. For example the number of memory controllers 215 and the number of memories 203, 205 used can vary. The memory control function could also be integrated into the control element 214 instead of using a discrete memory controller 215. In some embodiments the first memory 203 and/or the second memory 205 may be totally or partially integrated with the TME 200, but in other embodiments the first memory 203 and second memory 205 can be separate from the TME 200, for instance when a TME 200 is retroactively implemented in a device or design already having memory that the TME 200 can use.

[0047] Although the apparatus embodiments have been described in terms of parts, modular blocks, and engines to facilitate description, one or more routines, subroutines, components, subcomponents, registers, processors, circuits, software subroutines, and/or

software objects, or any combination thereof, could be substituted for one or several of the parts, modular blocks, and/or engines.

Methods

[0048] Once the reader 206 has obtained the content from the key 216, the TME 200 can perform various functions using the content, for example the TME 200 can perform a key 216 existence search, a data entry retrieval using the key, a key 216 (and associated data entry) insertion, and key 216 (and associated data entry) deletion. Each of the four aforementioned operations will be discussed below.

Performing a Key Lookup/ Data Entry Retrieval

[0049] Performing a key 216 lookup and performing a data entry retrieval based on a retrieved key 216 are similar. For illustrative purposes, it will be assumed that some data entries (in locations 272, 273, 274) and some related pointers 224, 226, 228 are already present in the tables 203, 205, although initially, before any data insertions, the second table 205 would be empty of pointers and have all validity bits set to "invalid." Likewise, in an initial state, the first table 203 would have all its memory locations free, and in one embodiment, linked together in a linked-list.

[0050] **Fig. 7** is a flowchart of an example data retrieval method, according to one aspect of the invention. First, a key is read and/or hashed for a content 700. A pointer memory location corresponding to the content is addressed using the content 702. A validity bit in the pointer memory location is read to determine if a data entry associated with the key is present in a first location in a first memory 704, 706. If the validity bit indicates that a

data entry for the key is not present, the data entry retrieval ends 708. If the validity bit indicates that a data entry for the key is present, then a pointer stored in the pointer memory location is used to find the data entry in the first location 710. The method is particularly suitable for managing a VLAN rule table, in which case the key is a VLAN tag and each data entry is a VLAN rule.

[0051] A TME 200 may be used to perform the method described above.

[0052] In accordance with one aspect of the invention, when a key 216 is received by the TME 200 the second memory 205 is arranged and/or selected so that the logical and/or physical position of each memory location 238-262 corresponds to the content of the key 216. Each physical and/or logical position (e.g., 238-262) in the second memory 205 stores a pointer and a validity bit that correspond to the content of a possible key 216 that could be received. The content of the key describes or represents a physical and/or logical position in the table/memory. Thus, after the key 216 is read and/or hashed by the reader 206, the executive table engine 208 can proceed directly to the proper location in the second memory 205 using the key content as an address.

[0053] For both the key lookup and the data entry retrieval operations, the executive table engine 208 proceeds to the location in the second memory 205 indicated by the content of the key 216 and reads a validity bit stored at the given location to determine if a pointer directed to a data entry for the key 216 has been stored there. If only a key lookup is being performed and the validity bit is "true," that is, the validity bit indicates that a valid data entry for the key 216 is present, then the key lookup operation is complete and requires no further action. In other words, for key lookups, which test for the mere presence of the key 216, or a representation of the key, the operation does not have to proceed any further than reading the validity bit. A data entry retrieval operation, however, requires additional action.

[0054] For a data retrieval operation, once the validity bit in the location corresponding to the content of the key 216 indicates the presence of a pointer for the key 216, then the pointer is followed to a data entry for the key in a data entry location in the first memory 203. For example, if the key 216 content is "4," the executive table engine 208 proceeds to memory location "4" 246 of the second memory 205 and reads the validity bit stored at memory location "4" 246 which, in the illustrated example is set to true ("1") indicating the presence of a valid pointer 224 for the key 216. The pointer 224 directs the executive table engine 208 to the data entry stored at the memory location 272 of the first memory 203. The executive table engine 208 can then retrieve the data entry.

[0055] Although in this embodiment a validity bit value of "1" indicates the existence of a data entry for the key 216 in the first memory 203 and a "0" indicates the absence of a data entry, in other embodiments the inverse may well be true, where "0" is used to indicate validity and "1" used to indicate invalidity.

Performing a Key Insertion or Deletion

[0056] The TME 200 can perform data entry 290 insertion or data entry 290 deletion operations in addition to the key lookup and data entry retrieval operations described above. Although a TME 200 can be used with a static table of data entries, the insertion and deletion operations may be used in many types of applications that require a table of dynamically changing data entries, not just a static table with a fixed number of data entries.

[0057] **Fig. 8** is a flowchart of an example method for performing a "key insertion," according to one aspect of the invention, that is, inserting a data entry for the key and setting a pointer to the data entry in a memory location representing the key. Thus, the

key insertion method is a method for building a data table. A data entry associated with a key is inserted in a first location of a first memory to begin building a table of data entries 800. The data entry may be inserted by an executive table engine 208 of a TME 200. Specifically, the data entry inserter/deleter 408 of the TME 200 can be used to perform the insertion. A pointer to the data entry is inserted in a second location in a second memory to begin building a table of pointers 802. The second location is selected so that a content of the key gives the address and/or position of the second location directly without a search through a list of keys or other entries. Hence, the address and/or position of the second location represents a content of the key. A pointer engine 404 component of the executive table engine 208 may be used to perform the pointer insertion. A validity bit in the second location is set to indicate the presence of the data entry associated with the key 804. The pointer engine 404 may also be used to set the validity bit.

[0058] Using a key with a content of "4" as an example for data entry 290 insertion, the TME 200 first performs the key 216 lookup operation discussed above and reads the validity bit in the memory location 246 representing the key 216 to determine whether a data entry is currently stored for the key 216. Once it has determined that no data entry is already stored for the key 216, the TME 200 receives the data entry 290 to be inserted and the executive table engine 208 stores the data entry 290 in the first available free memory location 272 in the first memory 203. If a list is being kept of free memory locations for data entries, then the memory location used by the inserted data entry is deleted from the list and the free list engine 210 reestablishes a new first available free location. The executive table engine 208 then places a pointer 224 pointing to the stored data entry into the memory location 246 in the second memory 205; the memory location 246 corresponding to the content "4" of the key 216. Since the memory location 272 receiving the data entry 290 has an address of "0000000000," the pointer 224 consists of address

"0000000000." Finally, the executive table engine 208 sets the validity bit for the memory location 246 where the new pointer resides to "valid," indicating that a valid data entry has been placed for the key 216 with content "4."

[0059] In this embodiment, the validity bit is set last in case an error occurs during the operation, so that the value of the validity bit gives as accurate an indication as possible of the presence or absence of a data entry for a given key. An error will result in the validity bit remaining in an "invalid" state, indicating no data entry for the key 216.

[0060] In one embodiment, the data entry deletion operation follows a sequence similar to the insertion sequence, except that a data entry and a pointer are removed instead of inserted.

[0061] **Fig. 9** is a flowchart of an example data deletion method, according to one aspect of the invention. The data entry associated with a key is deleted from a data entry location in a first memory 900. Pointers in a linked-list of free data entry locations are adjusted to include the data entry location freed by the data entry being deleted 902. A pointer to the data entry just removed is deleted from a pointer location in a second memory, wherein the pointer location represents the content of the key 904. Then, a validity bit in the pointer location is set to indicate the absence of a data entry associated with the key 906.

[0062] A TME 200 may be used to perform the method for deleting a data entry. Using a key 216 with content "4" as an example, the data entry residing at memory location 272 is deleted by the executive table engine 208. The pointer 224 from the memory location 246 in the second memory 205 is also removed. The newly freed memory location 272 in the first memory 203 is reintegrated into the list of free data entry memory locations (e.g., memory locations 275-276 and others in the table 203 that are empty). Lastly, the validity bit at the memory location 246 is set to indicate the absence of a data entry for the particular key 216. The reintegration of the freed memory location 272 into the list of free

data entry memory locations 275-276 may vary in its timing relative to the deletion of a data entry and a pointer. However, in this embodiment, the reintegration of the freed memory space (272 if the data entry there is being deleted) into the list of free data entry memory locations 275-276 is carried out dynamically as pointers to and from the data entry being deleted are rearranged, and is carried out by the free list engine 210.

Alternate Embodiment

[0063] **Fig. 10** is a graphical representation of an article of manufacture comprising a machine-readable medium 1000 having content 1002, that causes a host device to implement one or more embodiments, aspects, and/or methods of a table management engine of the invention. The content may be instructions, such as computer instructions, or may be design information allowing implementation. The content causes a machine to implement a method and/or apparatus of the invention, including inserting a data entry associated with a key in a data entry location 272 of a first memory 203, inserting a pointer 224 to the data entry in a pointer location in a second memory 205, wherein the address and/or position of the pointer location 246 represents a content of the key, and setting a validity bit in the pointer location 246 to indicate the presence of the data entry associated with the key in the data entry location 272.

[0064] The key received by the hosting machine may be a 12-bit VLAN tag. The hosting machine may implement a VLAN rule table having a 1K VLAN rule RAM, wherein each VLAN rule is the data entry for a key. When a VLAN rule is stored in the 1K rule RAM, a pointer to the VLAN rule is placed in a second RAM, specifically a 4K pointer RAM, at a location in the 4K pointer RAM representing the content of the key. In order to utilize the entire 1K VLAN rule RAM, the pointer is ten bits in length. A validity bit is also

stored at the pointer location in the 4K pointer RAM to indicate whether a valid VLAN rule is present for a given key content. Thus, the TME implemented by the machine addresses the 4K pointer RAM quickly using the content of a received VLAN tag, and quickly ascertains the presence or absence of a valid VLAN rule for the key by merely reading the validity bit. The high speed of the TME implemented by the machine is accomplished without the special hardware requirements and/or design complexity of a CAM chip.

[0065] The methods and apparatuses of the invention may be provided partially as a computer program product that may include the machine-readable medium. The machine-readable medium may include, but is not limited to, floppy diskettes, optical disks, CD-ROMs, magneto-optical disks, ROMs, RAMs, EPROMs, EEPROMs, magnetic or optical cards, flash memory, or other type of media suitable for storing electronic instructions. Moreover, parts may also be downloaded as a computer program product, wherein the program may be transferred from a remote computer to a requesting computer by way of data signals embodied in a carrier wave or other propagation media via a communication link (e.g., a modem or network connection). In this regard, the article of manufacture may well comprise such a carrier wave or other propagation media.

[0066] While the invention has been described in terms of several embodiments, those skilled in the art will recognize that the invention is not limited to the embodiments described but can be practiced with modification and alteration within the spirit and scope of the appended claims. The description is thus to be regarded as illustrative instead of limiting.

CLAIMS

What is claimed is:

1. An apparatus, comprising:

a reader to read and/or hash a key to determine a key content; and

an executive table engine to:

receive the key content from the reader,

store a data entry associated with the key in a data entry location in memory,

assign a pointer in a pointer location in memory, wherein the address of the pointer location is the key content and the pointer is directed to the data entry, and

assign a validity bit in the pointer location to indicate the presence of the data entry in the data entry location.

2. The apparatus of claim 1, wherein the executive table engine deletes the data entry from the data entry location, deletes the pointer from the pointer location, and assigns the validity bit to indicate the absence of the data entry.

3. The apparatus of claim 1, further comprising a free list engine to manage a list of free memory locations available for data entries and a head pointer register to maintain free memory location pointers within a linked-list of the free memory locations.

4. The apparatus of claim 1, the executive table engine further comprising:

a key indexer to find the pointer location using the key content and read the validity bit in the pointer location;

a data entry retriever to retrieve the data entry if the validity bit indicates the presence of the data entry associated with the key;

a data entry inserter/deleter to insert and/or delete the data entry in the data location;

and

a pointer engine to insert and/or delete the pointer and assign the validity bit associated with the pointer location.

5. A method, comprising:

inserting a data entry associated with a key in a data entry location in a first memory;

inserting a pointer to the data entry in a pointer location in a second memory, wherein the address and/or position of the pointer location represents a content of the key; and

setting a validity bit in the pointer location to indicate a presence of the data entry associated with the key.

6. The method of claim 5, wherein the depth of the first memory is based on a pointer length used in the second memory.

7. The method of claim 5, wherein empty data entry locations in the first memory are linked together in a list of free data entry locations.

8. The method of claim 7, further comprising pointing to a first available free data entry location in the list of free data entry locations using a head pointer register.

9. The method of claim 8, wherein the first available free data entry location for inserting the data entry is unlinked from the list of free data entry locations by updating the head pointer to point to a next available free data entry location.

10. The method of claim 5, wherein a ten bit pointer length is used in the second memory and 1024 memory locations are used in the first memory.

11. The method of claim 10, wherein the key is a virtual local area network (VLAN) tag having content represented by a memory location in the second memory and the data entry is a VLAN rule.

12. A method, comprising:

reading and/or hashing a key for a content;

reading a validity bit in a pointer location in memory to determine if a data entry associated with the key is present in a data entry location in memory, wherein the pointer location represents the content of the key; and

following a pointer in the pointer location to the data entry in the data entry location if the validity bit indicates the presence of the data entry in memory.

13. The method of claim 12, wherein the key is a VLAN tag and the data entry is a VLAN rule.

14. A method, comprising:

inserting a data entry associated with a key into a free data entry location in a data entry memory;

adjusting one or more pointers to exclude the free data entry location from a linked-list of free data entry locations;

inserting a pointer into a pointer location in memory, wherein the pointer is directed to the data entry and the address and/or position of the pointer location in memory represents the content of the key; and

setting a validity bit in the pointer location to indicate the presence of the data entry associated with the key.

15. The method of claim 14, wherein the key is a VLAN tag and the data entry is a VLAN rule.

16. A method, comprising:

deleting a data entry associated with a key from a data entry location in a data entry memory;

adjusting one or more free memory location pointers to include the data entry location in a linked-list of free data entry locations;

deleting a data entry pointer from a data entry pointer location in memory, wherein the data entry pointer is directed to the data entry location and the address and/or position of the data entry pointer location in memory represents the content of the key; and

setting a validity bit in the data entry pointer location to indicate the absence of the data entry associated with the key.

17. The method of claim 16, wherein the key is a VLAN tag and the data entry is a VLAN rule.

18. An article of manufacture, comprising:

a machine-readable medium containing content that, when executed, causes an accessing machine to:

read and/or hash a key to determine a key content;

store a data entry associated with the key in a data location in a first memory;

store a pointer and a validity bit in a pointer location in a second memory, wherein the pointer location represents the key content and the validity bit indicates the presence of the pointer in the pointer location; and

assign and/or delete the pointer based on the key content, and manage the validity bit to indicate the presence and/or absence of the pointer in the pointer location.

19. The article of manufacture of claim 18, wherein the content causes the accessing machine to read the validity bit in a pointer location representing the key and retrieve the data entry if the validity bit indicates the presence of a data entry associated with the key.

20. The article of manufacture of claim 18 wherein the content causes the accessing machine to manage free locations available for a data entry in the first memory and maintain pointers within a linked-list of the free locations.

21. A computing system, comprising:

a table management engine to:

store a data entry associated with a key in a data entry location in memory,

assign a pointer to a pointer location in memory, wherein the address of the pointer location corresponds to content from the key and the pointer is directed to the data entry,

assign a validity bit in the pointer location to indicate the presence of the data entry in the data entry location; and

at least one memory to store the data entry, the pointer, and the validity bit.

22. The computing system of claim 21, further comprising:

a reader to read and/or hash the key to determine the content;

a key indexer to find the pointer location based on the content and read the validity bit in the pointer location;

a data entry retriever to retrieve the data entry if the validity bit indicates the presence of the data entry associated with the key;

a data entry inserter/deleter to insert and/or delete the data entry in the data location;

and

a pointer engine to insert and/or delete the pointer and assign the validity bit associated with the pointer location.

23. The computing system of claim 21, further comprising a free list engine to manage a list of free memory locations available for the data entries including a head pointer register to maintain free memory location pointers within the list of the free memory locations.

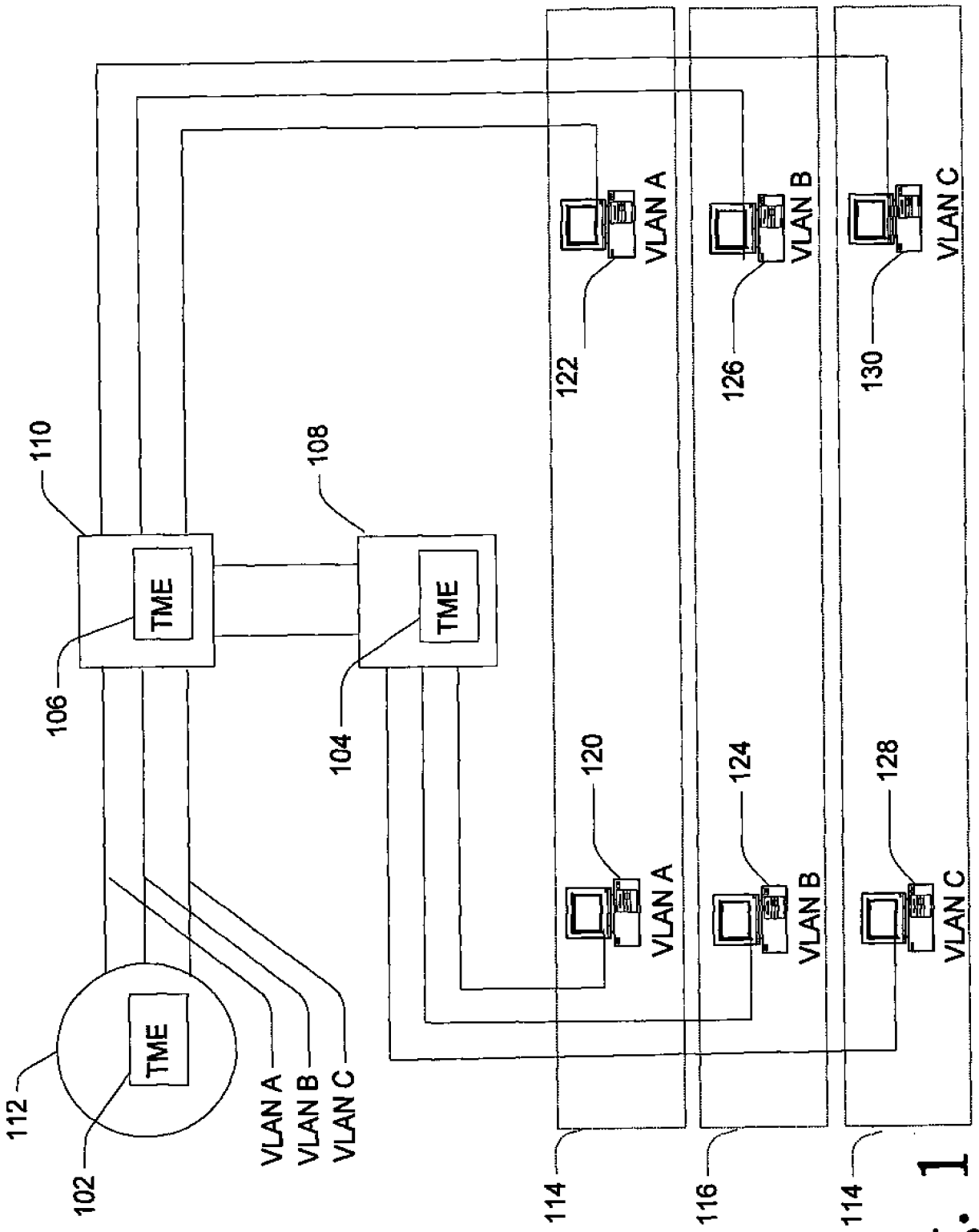
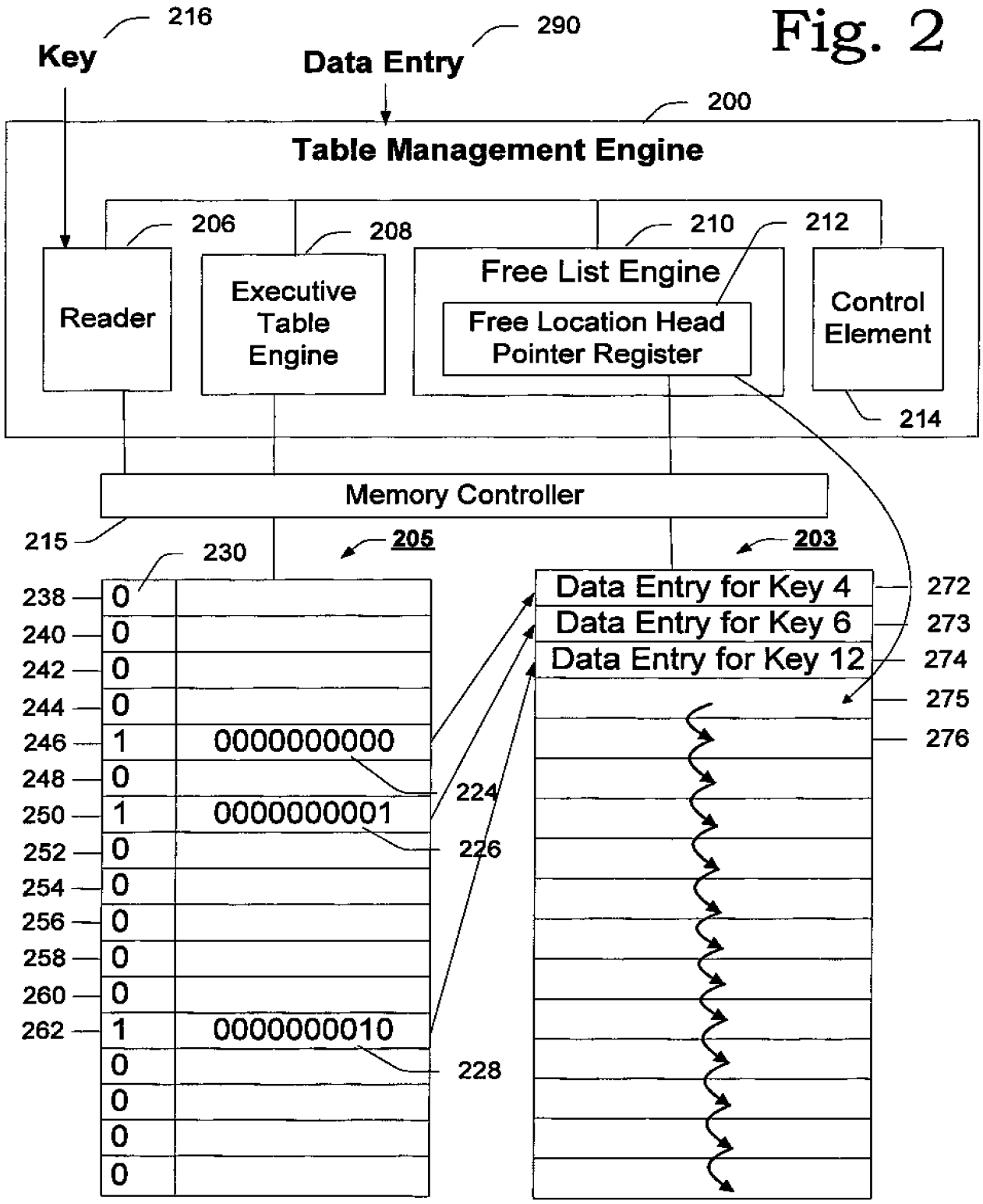

TME
TME
TME
110
108
106
104
112
102
VLAN A
VLAN B
VLAN C
122
126
130
120
124
128
VLAN A
VLAN B
VLAN C
VLAN A
VLAN B
VLAN C
114
116
114

Fig. 1

Fig. 2


Key
216
Data Entry
290
200
Table Management Engine
206
208
210
212
Reader
Executive Table Engine
Free List Engine
Free Location Head Pointer Register
Control Element
214
Memory Controller
215
205
203
230
238
240
242
244
246
248
250
252
254
256
258
260
262
0000000000
0000000001
0000000010
224
226
228
Data Entry for Key 4
Data Entry for Key 6
Data Entry for Key 12
272
273
274
275
276

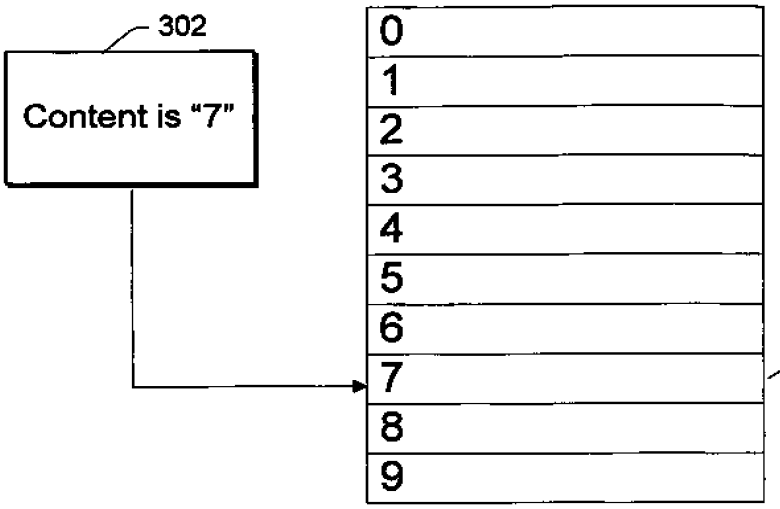


Fig. 3

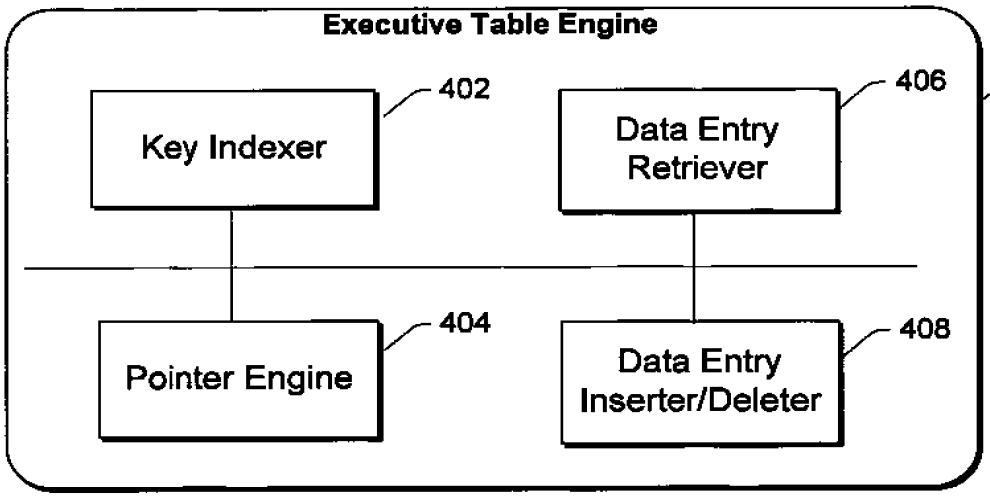


Fig. 4

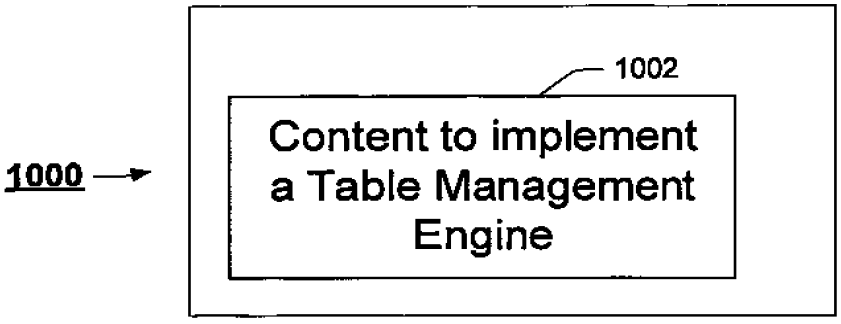


Fig. 10

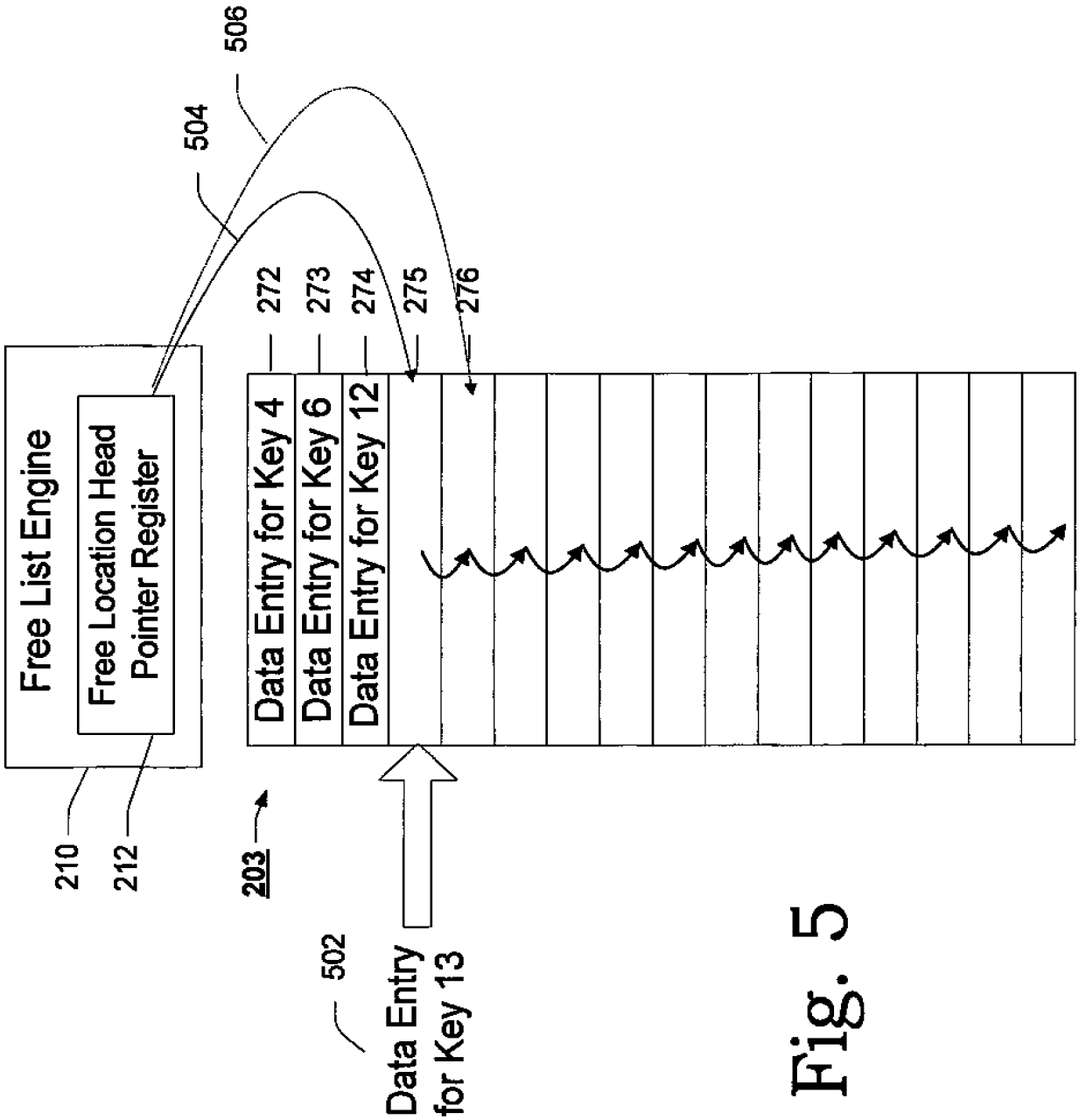

506
504
272
273
274
275
276
Free List Engine
Free Location Head Pointer Register
Data Entry for Key 4
Data Entry for Key 6
Data Entry for Key 12
210
212
203
502
Data Entry for Key 13

Fig. 5

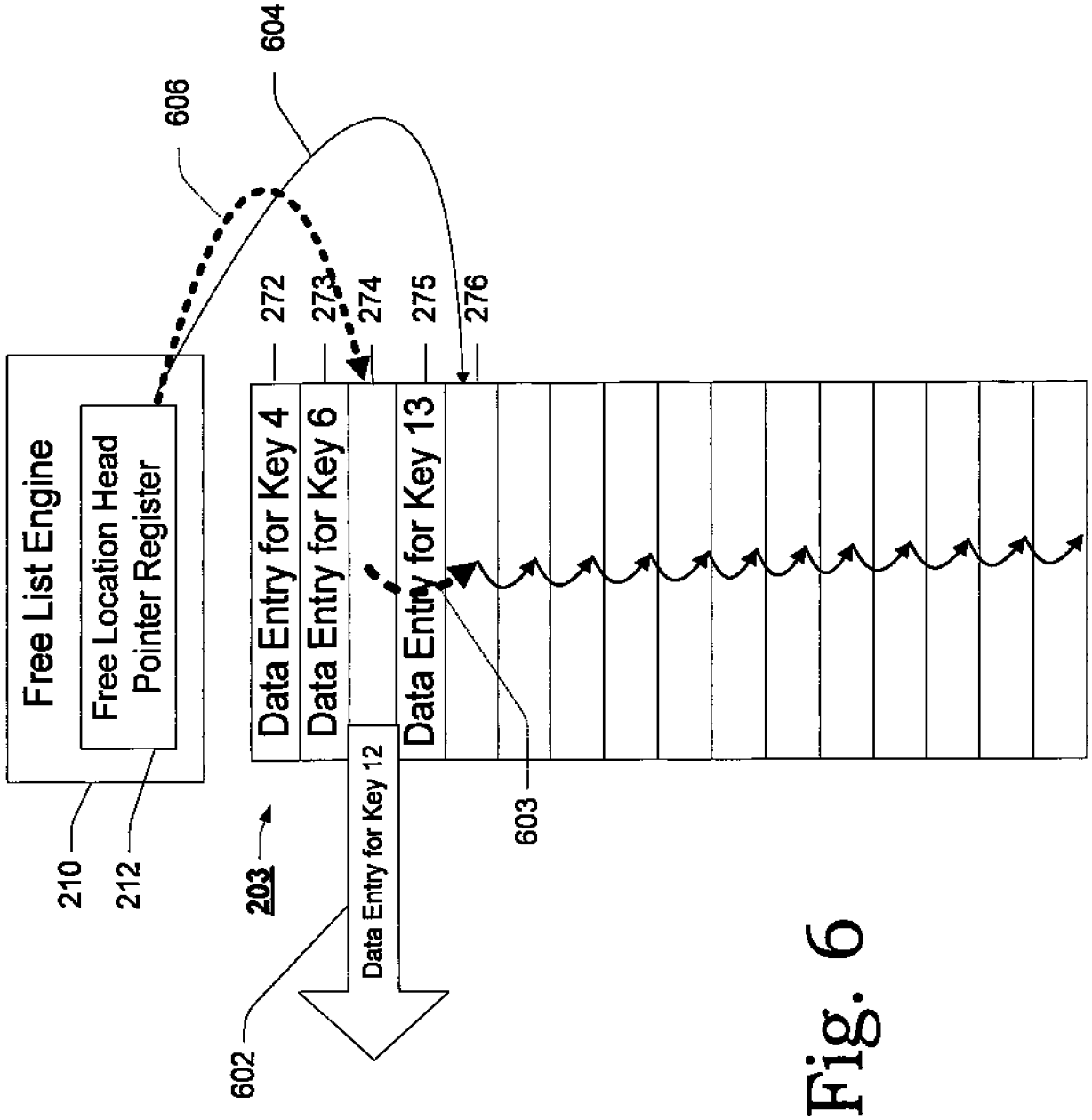

Free List Engine
Free Location Head Pointer Register
210
212
203
602
604
606
272
273
274
275
276
Data Entry for Key 4
Data Entry for Key 6
Data Entry for Key 13
Data Entry for Key 12
603

Fig. 6

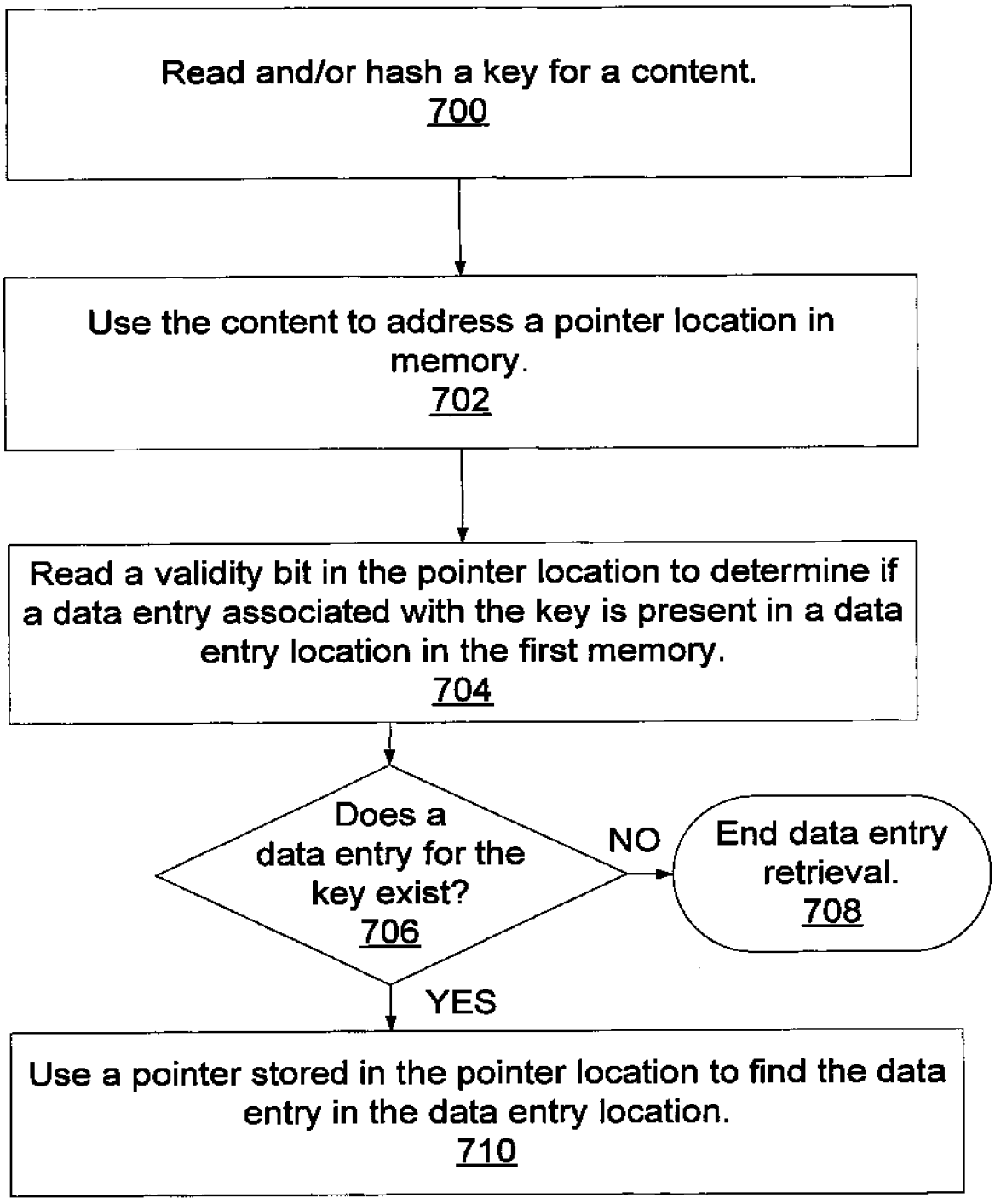

Read and/or hash a key for a content.
700
Use the content to address a pointer location in memory.
702
Read a validity bit in the pointer location to determine if a data entry associated with the key is present in a data entry location in the first memory.
704
Does a data entry for the key exist?
706
NO
End data entry retrieval.
708
YES
Use a pointer stored in the pointer location to find the data entry in the data entry location.
710

Fig. 7

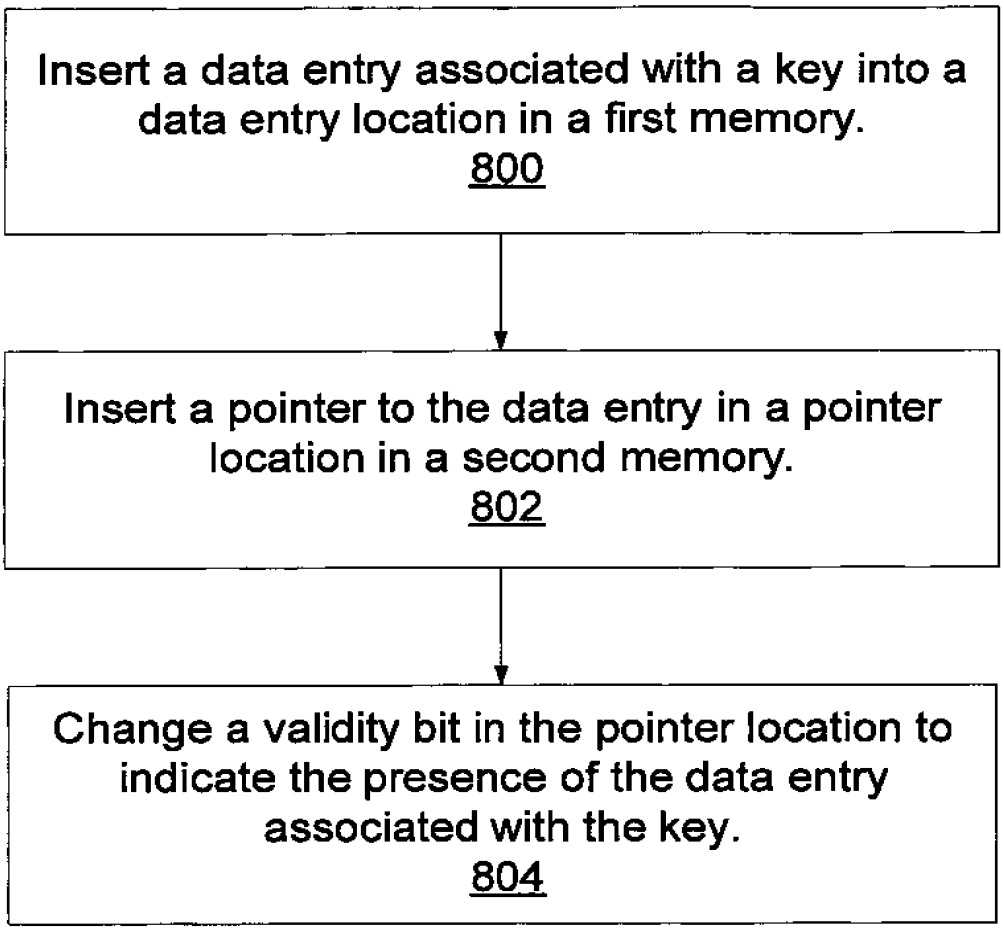

Insert a data entry associated with a key into a data entry location in a first memory.
800
Insert a pointer to the data entry in a pointer location in a second memory.
802
Change a validity bit in the pointer location to indicate the presence of the data entry associated with the key.
804

Fig. 8

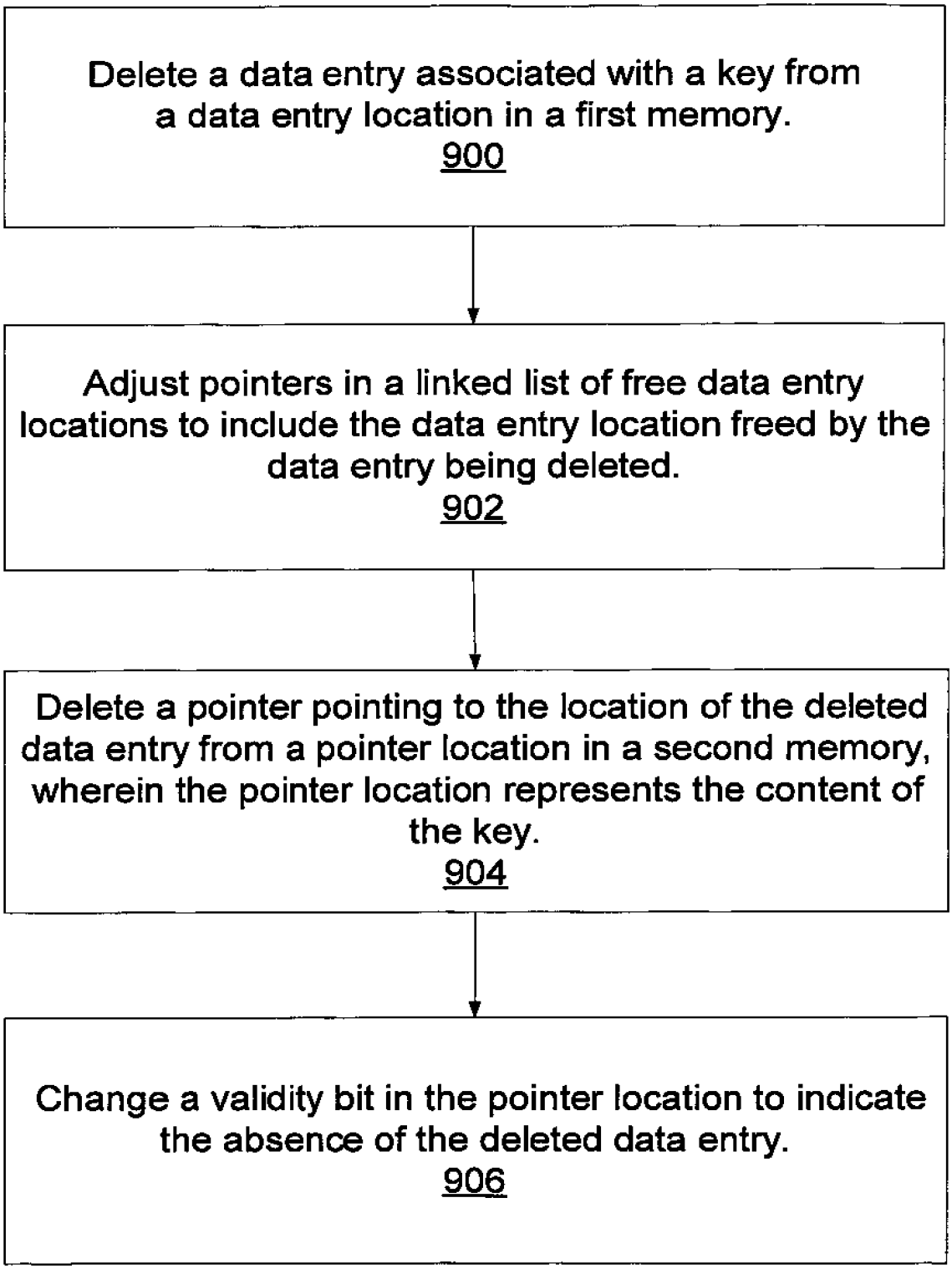

Delete a data entry associated with a key from a data entry location in a first memory.
900
Adjust pointers in a linked list of free data entry locations to include the data entry location freed by the data entry being deleted.
902
Delete a pointer pointing to the location of the deleted data entry from a pointer location in a second memory, wherein the pointer location represents the content of the key.
904
Change a validity bit in the pointer location to indicate the absence of the deleted data entry.
906

Fig. 9

INTERNATIONAL SEARCH REPORT

PCT/US 03/06641

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04L12/56 H04L12/46

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 H04L G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 6 034 958 A (WICKLUND GORAN) 7 March 2000 (2000-03-07) column 3, line 39 -column 4, line 67 column 5, line 14 - line 67 column 7, line 44 -column 8, line 41 figures 1-6	1-23
Y	US 6 091 707 A (RUNALDUE THOMAS J ET AL) 18 July 2000 (2000-07-18) column 5, line 41 -column 6, line 18 column 6, line 58 -column 8, line 3 column 9, line 39 -column 10, line 21 figures 2-5	1-23
A	EP 1 130 855 A (CIT ALCATEL) 5 September 2001 (2001-09-05) column 4, line 39 -column 7, line 44 figures 2-7	1-23

-/--

[X] Further documents are listed in the continuation of box C. [X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 August 2003	26/08/2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Lastoria, G

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

PCT/US 03/06641

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	YU D ET AL: "FORWARDING ENGINE FOR FAST ROUTING LOOKUPS AND UPDATES" 1999 IEEE GLOBAL TELECOMMUNICATIONS CONFERENCE. GLOBECOM'99. SEAMLESS INTERCONNECTION FOR UNIVERSAL SERVICES. RIO DE JANEIRO, BRAZIL, DEC. 5-9, 1999, IEEE GLOBAL TELECOMMUNICATIONS CONFERENCE, NEW YORK, NY: IEEE, US, vol. 2, 5 December 1999 (1999-12-05), pages 1556-1564, XP001016965 ISBN: 0-7803-5797-3 page 1559, left-hand column, line 1 - line 17 page 1561, right-hand column, line 31 -page 1562, left-hand column, line 23	1-23

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

PCT/US 03/06641

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6034958 A	07-03-2000	AU 8362798 A	08-02-1999
		CN 1272296 T	01-11-2000
		GB 2342250 A ,B	05-04-2000
		JP 2001509654 T	24-07-2001
		WO 9903298 A1	21-01-1999
US 6091707 A	18-07-2000	NONE	
EP 1130855 A	05-09-2001	CN 1311607 A	05-09-2001
		EP 1130855 A2	05-09-2001
		JP 2001285333 A	12-10-2001